

OIPE #2

CRF Errors Corrected by the STIC Systems Branch

Serial Number: 10/074,527

CRF Processing Date: 3/12/2002
Edited by: AC
Verified by: AC (STIC staff)

ENTERED

- ☐ Changed a file from non-ASCII to ASCII
- ☐ Changed the margins in cases where the sequence text was "wrapped" down to the next line.
- ☐ Edited a format error in the Current Application Data section, specifically: _____
- ☐ Edited the Current Application Data section with the actual current number. The number inputted by the applicant was ☐ the prior application data; or ☐ other _____
- ☐ Added the mandatory heading and subheadings for "Current Application Data".
- ☐ Edited the "Number of Sequences" field. The applicant spelled out a number instead of using an integer.
- ☐ Changed the spelling of a mandatory field (the headings or subheadings), specifically: _____
- ☐ Corrected the SEQ ID NO when obviously incorrect. The sequence numbers that were edited were: _____
- ☐ Inserted or corrected a nucleic number at the end of a nucleic line. SEQ ID NO's edited: _____
- ☐ Corrected subheading placement. All responses must be on the same line as each subheading. If the applicant placed a response below the subheading, this was moved to its appropriate place.
- ☐ Inserted colons after headings/subheadings. Headings edited included: _____
- ☐ Deleted extra, invalid, headings used by an applicant, specifically: _____
- ☒ Deleted: ☒ non-ASCII "garbage" at the beginning/end of files; ☐ secretary initials/filename at end of file; ☐ page numbers throughout text; ☐ other invalid text, such as _____
- ☐ Inserted mandatory headings, specifically: _____
- ☐ Corrected an obvious error in the response, specifically: _____
- ☐ Edited identifiers where upper case is used but lower case is required, or vice versa.
- ☐ Corrected an error in the Number of Sequences field, specifically: _____
- ☐ A "Hard Page Break" code was inserted by the applicant. All occurrences had to be deleted.
- ☐ Deleted *ending* stop codon in amino acid sequences and adjusted the "(A)Length:" field accordingly (error due to a PatentIn bug). Sequences corrected: _____
- ☐ Other: _____

*Examiner: The above corrections must be communicated to the applicant in the first Office Action. DO NOT send a copy of this form.

3/1/95



RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/074,527

DATE: 03/12/2002

TIME: 17:47:57

Input Set : A:\pto.amc.txt

Output Set: N:\CRF3\03122002\J074527.raw

26

```

4 <110> APPLICANT: Olandt, Peter J.
5     Meyers, Rachel E.
6     Galvin, Katherine A.
7     Millennium Pharmaceuticals Inc.
9 <120> TITLE OF INVENTION: 33945, A Human Glycosyltransferase and
10    Uses Therefor
12 <130> FILE REFERENCE: MPI2001-018P1RCPI(M)
C--> 14 <140> CURRENT APPLICATION NUMBER: US/10/074,527
C--> 14 <141> CURRENT FILING DATE: 2002-02-12
14 <150> PRIOR APPLICATION NUMBER: 60/269202
15 <151> PRIOR FILING DATE: 2001-02-15
17 <160> NUMBER OF SEQ ID NOS: 9
19 <170> SOFTWARE: FastSEQ for Windows Version 4.0
21 <210> SEQ ID NO: 1
22 <211> LENGTH: 2850
23 <212> TYPE: DNA
24 <213> ORGANISM: homo sapiens
26 <220> FEATURE:
27 <221> NAME/KEY: CDS
28 <222> LOCATION: (81)...(1826)
30 <400> SEQUENCE: 1
31 ccggctcggt accactataa cggccgccag tgtgctggaa ttgcccttg cgcagatcgc 60
32 tggctgcagt tggcggggcgc atg tgg ggg cgc acg gcg cgg cgc tgc ccg 113
33                               Met Trp Gly Arg Thr Ala Arg Arg Arg Cys Pro
34                               1           5           10
36 cgg gaa ctg cgg cgc ggc cgg gag gcg ctg ttg gtg ctc ctg gcg cta 161
37 Arg Glu Leu Arg Arg Gly Arg Glu Ala Leu Leu Val Leu Leu Ala Leu
38                               15           20           25
40 ctg gcg ttg gcc ggg ctg ggc tcg gtg ctg cgg gcg cag cgt ggg gcc 209
41 Leu Ala Leu Ala Gly Leu Gly Ser Val Leu Arg Ala Gln Arg Gly Ala
42                               30           35           40
44 ggg gcc ggg gct gcc gag ccg gga ccc ccg cgc acc ccg cgc ccc ggg 257
45 Gly Ala Gly Ala Ala Glu Pro Gly Pro Pro Arg Thr Pro Arg Pro Gly
46                               45           50           55
48 cgg cgc gag ccg gtc atg ccg cgg ccg ccg gtg ccg gcg aac gcg ctg 305
49 Arg Arg Glu Pro Val Met Pro Arg Pro Pro Val Pro Ala Asn Ala Leu
50 60           65           70           75
52 ggc gcg cgg ggc gag gcg gtg cgg ctg cag ctg cag ggc gag gag ctg 353
53 Gly Ala Arg Gly Glu Ala Val Arg Leu Gln Leu Gln Gly Glu Glu Leu
54                               80           85           90
56 cgg ctg cag gag gag agc gtg cgg ctg cac cag att aac atc tac ctc 401
57 Arg Leu Gln Glu Glu Ser Val Arg Leu His Gln Ile Asn Ile Tyr Leu
58                               95           100          105

```

RAW SEQUENCE LISTING

DATE: 03/12/2002

PATENT APPLICATION: US/10/074,527

TIME: 17:47:57

Input Set : A:\pto.amc.txt

Output Set: N:\CRF3\03122002\J074527.raw

```

60 agc gac cgc atc tca ctg cac cgc cgc ctg ccc gag cgc tgg aac ccg 449
61 Ser Asp Arg Ile Ser Leu His Arg Arg Leu Pro Glu Arg Trp Asn Pro
62 110 115 120
64 ctg tgc aaa gag aag aaa tat gat tat gat aat ttg ccc agg aca tct 497
65 Leu Cys Lys Glu Lys Lys Tyr Asp Tyr Asp Asn Leu Pro Arg Thr Ser
66 125 130 135
68 gtt atc ata gca ttt tat aat gaa gcc tgg tca act ctc ctt cgg aca 545
69 Val Ile Ile Ala Phe Tyr Asn Glu Ala Trp Ser Thr Leu Leu Arg Thr
70 140 145 150 155
72 gtt tac agt gtc ctt gag aca tcc ccg gat atc ctg cta gaa gaa gtg 593
73 Val Tyr Ser Val Leu Glu Thr Ser Pro Asp Ile Leu Leu Glu Glu Val
74 160 165 170
76 atc ctt gta gat gac tac agt gat aga gag cac ctg aag gag cgc ttg 641
77 Ile Leu Val Asp Asp Tyr Ser Asp Arg Glu His Leu Lys Glu Arg Leu
78 175 180 185
80 gcc aat gag ctt tcg gga ctg ccc aag gtg cgc ctg atc cgc gcc aac 689
81 Ala Asn Glu Leu Ser Gly Leu Pro Lys Val Arg Leu Ile Arg Ala Asn
82 190 195 200
84 aag aga gag ggc ctg gtg cga gcc ccg ctg ctg ggg gcg tct gcg gcg 737
85 Lys Arg Glu Gly Leu Val Arg Ala Arg Leu Leu Gly Ala Ser Ala Ala
86 205 210 215
88 agg ggc gat gtt ctg acc ttc ctg gac tgt cac tgt gag tgc cac gaa 785
89 Arg Gly Asp Val Leu Thr Phe Leu Asp Cys His Cys Glu Cys His Glu
90 220 225 230 235
92 ggg tgg ctg gag ccg ctg ctg cag agg atc cat gaa gag gag tcg gca 833
93 Gly Trp Leu Glu Pro Leu Leu Gln Arg Ile His Glu Glu Glu Ser Ala
94 240 245 250
96 gtg gtg tgc ccg gtg att gat gtg atc gac tgg aac acc ttc gaa tac 881
97 Val Val Cys Pro Val Ile Asp Val Ile Asp Trp Asn Thr Phe Glu Tyr
98 255 260 265
100 ctg ggg aac tcc ggg gag ccc cag atc ggc ggt ttc gac tgg agg ctg 929
101 Leu Gly Asn Ser Gly Glu Pro Gln Ile Gly Gly Phe Asp Trp Arg Leu
102 270 275 280
104 gtg ttc acg tgg cac aca gtt cct gag agg gag agg ata cgg atg caa 977
105 Val Phe Thr Trp His Thr Val Pro Glu Arg Glu Arg Ile Arg Met Gln
106 285 290 295
108 tcc ccc gtc gat gtc atc agg tct cca aca atg gct ggt ggg ctg ttt 1025
109 Ser Pro Val Asp Val Ile Arg Ser Pro Thr Met Ala Gly Gly Leu Phe
110 300 305 310 315
112 gct gtg agt aag aaa tat ttt gaa tat ctg ggg tct tat gat aca gga 1073
113 Ala Val Ser Lys Lys Tyr Phe Glu Tyr Leu Gly Ser Tyr Asp Thr Gly
114 320 325 330
116 atg gaa gtt tgg gga gga gaa aac ctc gaa ttt tcc ttt agg atc tgg 1121
117 Met Glu Val Trp Gly Gly Glu Asn Leu Glu Phe Ser Phe Arg Ile Trp
118 335 340 345
120 cag tgt ggt ggg gtt ctg gaa aca cac cca tgt tcc cat gtt ggc cat 1169
121 Gln Cys Gly Gly Val Leu Glu Thr His Pro Cys Ser His Val Gly His
122 350 355 360
124 gtt ttc ccc aag caa gct ccc tac tcc cgc aac aag gct ctg gcc aac 1217

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/074,527

DATE: 03/12/2002

TIME: 17:47:57

Input Set : A:\pto.amc.txt

Output Set: N:\CRF3\03122002\J074527.raw

```

125 Val Phe Pro Lys Gln Ala Pro Tyr Ser Arg Asn Lys Ala Leu Ala Asn
126      365      370      375
128 agt gtt cgt gca gct gaa gta tgg atg gat gaa ttt aaa gag ctc tac 1265
129 Ser Val Arg Ala Ala Glu Val Trp Met Asp Glu Phe Lys Glu Leu Tyr
130 380      385      390      395
132 tac cat cgc aac ccc cgt gcc cgc ttg gaa cct ttt ggg gat gtg aca 1313
133 Tyr His Arg Asn Pro Arg Ala Arg Leu Glu Pro Phe Gly Asp Val Thr
134      400      405      410
136 gag agg aag cag ctc cgg gac aag ctc cag tgt aaa gac ttc aag tgg 1361
137 Glu Arg Lys Gln Leu Arg Asp Lys Leu Gln Cys Lys Asp Phe Lys Trp
138      415      420      425
140 ttc ttg gag act gtg tat cca gaa ctg cat gtg cct gag gac agg cct 1409
141 Phe Leu Glu Thr Val Tyr Pro Glu Leu His Val Pro Glu Asp Arg Pro
142      430      435      440
144 ggc ttc ttc ggg atg ctc cag aac aaa gga cta aca gac tac tgc ttt 1457
145 Gly Phe Phe Gly Met Leu Gln Asn Lys Gly Leu Thr Asp Tyr Cys Phe
146      445      450      455
148 gac tat aac cct ccc gat gaa aac cag att gtg gga cac cag gtc att 1505
149 Asp Tyr Asn Pro Pro Asp Glu Asn Gln Ile Val Gly His Gln Val Ile
150 460      465      470      475
152 ctg tac ctc tgt cat ggg atg ggc cag aat cag ttt ttc gag tac acg 1553
153 Leu Tyr Leu Cys His Gly Met Gly Gln Asn Gln Phe Phe Glu Tyr Thr
154      480      485      490
156 tcc cag aaa gaa ata cgc tat aac acc cac cag cct gag ggc tgc att 1601
157 Ser Gln Lys Glu Ile Arg Tyr Asn Thr His Gln Pro Glu Gly Cys Ile
158      495      500      505
160 gct gtg gaa gca gga atg gat acc ctt atc atg cat ctc tgc gaa gaa 1649
161 Ala Val Glu Ala Gly Met Asp Thr Leu Ile Met His Leu Cys Glu Glu
162      510      515      520
164 act gcc cca gag aat cag aag ttc atc ttg cag gag gat gga tct tta 1697
165 Thr Ala Pro Glu Asn Gln Lys Phe Ile Leu Gln Glu Asp Gly Ser Leu
166      525      530      535
168 ttt cac gaa cag tcc aag aaa tgt gtc cag gct gcg agg aag gag tcg 1745
169 Phe His Glu Gln Ser Lys Lys Cys Val Gln Ala Ala Arg Lys Glu Ser
170 540      545      550      555
172 agt gac agt ttc gtt cca ctc tta cga gac tgc acc aac tcg gat cat 1793
173 Ser Asp Ser Phe Val Pro Leu Leu Arg Asp Cys Thr Asn Ser Asp His
174      560      565      570
176 cag aaa tgg ttc ttc aaa gag cgc atg tta tga agcctcgtgt atcaaggagc 1846
177 Gln Lys Trp Phe Phe Lys Glu Arg Met Leu *
178      575      580
180 ccacgaagg agactgtgga gccaggactc tgcccaacaa agacttagct aagcagtgc 1906
181 cagaaccac caaaaactag gctgcattgc tttgaagagg caatcatttt gccatttgtg 1966
182 aaagtgtgt tggatttagt aaaaatgtga ataagctttg tacttatttt gagaactttt 2026
183 taaatgttcc aaaataccct attttcaaag ggtaatcgta agatgttaac ctttggtatt 2086
184 tagaaaatta aaacottata atatttttct atcaarawrw awattttaca gtcgtgcctt 2146
185 ttactctcat tagcaaaaaa gataaagatt ttatttttgt atttacaaga attccaggt 2206
186 acgaagatat ctgcatgggt ggaaatcagg ttcaagcaac gtactttgca ttaactgata 2266
187 atacctcagc tgcggggtta aagttttccc agtatagaga gaactgtcact aggaacattg 2326

```

RAW SEQUENCE LISTING

DATE: 03/12/2002

PATENT APPLICATION: US/10/074,527

TIME: 17:47:57

Input Set : A:\pto.amc.txt

Output Set: N:\CRF3\03122002\J074527.raw

```

188 tattgattta ttcagggtcat tgagatcttc tagatgtatt ttaaaaagaa tgctttttgg 2386
189 ttatgtgttg ctaccacagt taacactcca taatgttcat gtcagccaaa gaggactaac 2446
190 caaagctgaa atctcagaga acaattttgct ttactaagct gagtcaactt gagagcgaac 2506
191 ttctaacaat gccgcactgt agtgtggctg gttctaccac tatgacttta aaacatgttt 2566
192 atatcatttt taatttttat gatacggtag tgtcaggagg aaatgtaatg ttctatatga 2626
193 aattcctttt tcaagtttgt tcattaataa cagttattaa tttaaatcag cgtttagagtt 2686
194 tgtgctgctg caactgctgt gaaaatttct ctgagtaatt ctgatttgtg aatgatccca 2746
195 gaccaaccct gagattttgt caacctgatt aagtcaatat gaatgattaa aaagatgtga 2806
196 gaacaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaa 2850
198 <210> SEQ ID NO: 2
199 <211> LENGTH: 581
200 <212> TYPE: PRT
201 <213> ORGANISM: homo sapiens
203 <400> SEQUENCE: 2
204 Met Trp Gly Arg Thr Ala Arg Arg Arg Cys Pro Arg Glu Leu Arg Arg
205 1 5 10 15
206 Gly Arg Glu Ala Leu Leu Val Leu Leu Ala Leu Leu Ala Leu Ala Gly
207 20 25 30
208 Leu Gly Ser Val Leu Arg Ala Gln Arg Gly Ala Gly Ala Gly Ala Ala
209 35 40 45
210 Glu Pro Gly Pro Pro Arg Thr Pro Arg Pro Gly Arg Arg Glu Pro Val
211 50 55 60
212 Met Pro Arg Pro Pro Val Pro Ala Asn Ala Leu Gly Ala Arg Gly Glu
213 65 70 75 80
214 Ala Val Arg Leu Gln Leu Gln Gly Glu Glu Leu Arg Leu Gln Glu Glu
215 85 90 95
216 Ser Val Arg Leu His Gln Ile Asn Ile Tyr Leu Ser Asp Arg Ile Ser
217 100 105 110
218 Leu His Arg Arg Leu Pro Glu Arg Trp Asn Pro Leu Cys Lys Glu Lys
219 115 120 125
220 Lys Tyr Asp Tyr Asp Asn Leu Pro Arg Thr Ser Val Ile Ile Ala Phe
221 130 135 140
222 Tyr Asn Glu Ala Trp Ser Thr Leu Leu Arg Thr Val Tyr Ser Val Leu
223 145 150 155 160
224 Glu Thr Ser Pro Asp Ile Leu Leu Glu Glu Val Ile Leu Val Asp Asp
225 165 170 175
226 Tyr Ser Asp Arg Glu His Leu Lys Glu Arg Leu Ala Asn Glu Leu Ser
227 180 185 190
228 Gly Leu Pro Lys Val Arg Leu Ile Arg Ala Asn Lys Arg Glu Gly Leu
229 195 200 205
230 Val Arg Ala Arg Leu Leu Gly Ala Ser Ala Ala Arg Gly Asp Val Leu
231 210 215 220
232 Thr Phe Leu Asp Cys His Cys Glu Cys His Glu Gly Trp Leu Glu Pro
233 225 230 235 240
234 Leu Leu Gln Arg Ile His Glu Glu Glu Ser Ala Val Val Cys Pro Val
235 245 250 255
236 Ile Asp Val Ile Asp Trp Asn Thr Phe Glu Tyr Leu Gly Asn Ser Gly
237 260 265 270
238 Glu Pro Gln Ile Gly Gly Phe Asp Trp Arg Leu Val Phe Thr Trp His

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/074,527

DATE: 03/12/2002

TIME: 17:47:57

Input Set : A:\pto.amc.txt

Output Set: N:\CRF3\03122002\J074527.raw

```

239          275          280          285
240 Thr Val Pro Glu Arg Glu Arg Ile Arg Met Gln Ser Pro Val Asp Val
241          290          295          300
242 Ile Arg Ser Pro Thr Met Ala Gly Gly Leu Phe Ala Val Ser Lys Lys
243 305          310          315          320
244 Tyr Phe Glu Tyr Leu Gly Ser Tyr Asp Thr Gly Met Glu Val Trp Gly
245          325          330          335
246 Gly Glu Asn Leu Glu Phe Ser Phe Arg Ile Trp Gln Cys Gly Gly Val
247          340          345          350
248 Leu Glu Thr His Pro Cys Ser His Val Gly His Val Phe Pro Lys Gln
249          355          360          365
250 Ala Pro Tyr Ser Arg Asn Lys Ala Leu Ala Asn Ser Val Arg Ala Ala
251          370          375          380
252 Glu Val Trp Met Asp Glu Phe Lys Glu Leu Tyr Tyr His Arg Asn Pro
253 385          390          395          400
254 Arg Ala Arg Leu Glu Pro Phe Gly Asp Val Thr Glu Arg Lys Gln Leu
255          405          410          415
256 Arg Asp Lys Leu Gln Cys Lys Asp Phe Lys Trp Phe Leu Glu Thr Val
257          420          425          430
258 Tyr Pro Glu Leu His Val Pro Glu Asp Arg Pro Gly Phe Phe Gly Met
259          435          440          445
260 Leu Gln Asn Lys Gly Leu Thr Asp Tyr Cys Phe Asp Tyr Asn Pro Pro
261          450          455          460
262 Asp Glu Asn Gln Ile Val Gly His Gln Val Ile Leu Tyr Leu Cys His
263 465          470          475          480
264 Gly Met Gly Gln Asn Gln Phe Phe Glu Tyr Thr Ser Gln Lys Glu Ile
265          485          490          495
266 Arg Tyr Asn Thr His Gln Pro Glu Gly Cys Ile Ala Val Glu Ala Gly
267          500          505          510
268 Met Asp Thr Leu Ile Met His Leu Cys Glu Glu Thr Ala Pro Glu Asn
269          515          520          525
270 Gln Lys Phe Ile Leu Gln Glu Asp Gly Ser Leu Phe His Glu Gln Ser
271          530          535          540
272 Lys Lys Cys Val Gln Ala Ala Arg Lys Glu Ser Ser Asp Ser Phe Val
273 545          550          555          560
274 Pro Leu Leu Arg Asp Cys Thr Asn Ser Asp His Gln Lys Trp Phe Phe
275          565          570          575
276 Lys Glu Arg Met Leu
277          580
280 <210> SEQ ID NO: 3
281 <211> LENGTH: 1746
282 <212> TYPE: DNA
283 <213> ORGANISM: homo sapiens
285 <220> FEATURE:
286 <221> NAME/KEY: CDS
287 <222> LOCATION: (1)...(1746)
289 <400> SEQUENCE: 3
290 atg tgg ggg cgc acg gcg cgg cgg cgc tgc ccg cgg gaa ctg cgg cgc 48
291 Met Trp Gly Arg Thr Ala Arg Arg Arg Cys Pro Arg Glu Leu Arg Arg

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/074,527

DATE: 03/12/2002
TIME: 17:47:58

Input Set : A:\pto.amc.txt
Output Set: N:\CRF3\03122002\J074527.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:9; Xaa Pos. 2,3,4,5,6,7,9,10,11,12,13,14,16,17,18,19,20,21



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/074,527

DATE: 02/28/2002

TIME: 10:02:19

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\02282002\J074527.raw

Does Not Comply
Corrected Diskette Needed

est

4 <110> APPLICANT: Olandt, Peter J.
5 Meyers, Rachel E.
6 Galvin, Katherine A.
7 Millennium Pharmaceuticals Inc.
9 <120> TITLE OF INVENTION: 33945, A Human Glycosyltransferase and
10 Uses Therefor
12 <130> FILE REFERENCE: MPI2001-018P1RCPI(M)
14 <140> CURRENT APPLICATION NUMBER: US/10/074,527
14 <141> CURRENT FILING DATE: 2002-02-12
14 <150> PRIOR APPLICATION NUMBER: 60/269202
15 <151> PRIOR FILING DATE: 2001-02-15
17 <160> NUMBER OF SEQ ID NOS: 9
19 <170> SOFTWARE: FastSEQ for Windows Version 4.0

ERRORED SEQUENCES

649 <210> SEQ ID NO: 9
650 <211> LENGTH: 22
651 <212> TYPE: PRT
652 <213> ORGANISM: Artificial Sequence
654 <220> FEATURE:
655 <223> OTHER INFORMATION: consensus
657 <221> NAME/KEY: VARIANT
658 <222> LOCATION: (1)...(22)
659 <223> OTHER INFORMATION: Xaa = any amino acid
661 <400> SEQUENCE: 9
W--> 662 Leu Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa Xaa Leu Xaa
663 1 5 10 15
W--> 664 Xaa Xaa Xaa Xaa Xaa Leu
665 20
E--> 667 - 1 -

delete

VERIFICATION SUMMARY

DATE: 02/28/2002

PATENT APPLICATION: US/10/074,527

TIME: 10:02:20

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\02282002\J074527.raw

L:14 M:270 C: Current Application Number differs, Replaced Current Application No

L:14 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:662 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9

L:664 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:9

L:667 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:9